NASA/TM—2003-212220

Threshold Assessment of Gear Diagnostic Tools on Flight and Test Rig Data

Paula J. Dempsey
Glenn Research Center, Cleveland, Ohio

Marianne Mosher and Edward M. Huff
Ames Research Center, Moffett Field, California

June 2003

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the Lead Center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at ***http://www.sti.nasa.gov***

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA Access Help Desk at 301–621–0134

- Telephone the NASA Access Help Desk at 301–621–0390

- Write to:
  NASA Access Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076

NASA/TM—2003-212220

Threshold Assessment of Gear Diagnostic Tools on Flight and Test Rig Data

Paula J. Dempsey
Glenn Research Center, Cleveland, Ohio

Marianne Mosher and Edward M. Huff
Ames Research Center, Moffett Field, California

Prepared for the
59th Annual Forum and Technology Display
sponsored by the American Helicopter Society
Phoenix, Arizona, May 6–8, 2003

National Aeronautics and
Space Administration

Glenn Research Center

June 2003

Available from

NASA Center for Aerospace Information
7121 Standard Drive
Hanover, MD 21076

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22100

Available electronically at http://gltrs.grc.nasa.gov

Threshold Assessment of Gear Diagnostic Tools on Flight and Test Rig Data

Paula J. Dempsey
NASA Glenn Research Center
Cleveland, Ohio

Marianne Mosher
NASA Ames Research Center
Moffett Field, CA

Edward M. Huff
NASA Ames Research Center
Moffett Field, CA

Abstract

A method for defining thresholds for vibration-based algorithms that provides the minimum number of false alarms while maintaining sensitivity to gear damage was developed. This analysis focused on two vibration based gear damage detection algorithms, FM4 and M8A. This method was developed using vibration data collected during surface fatigue tests performed in a spur gearbox rig. The thresholds were defined based on damage progression during tests with damage. The thresholds false alarm rates were then evaluated on spur gear tests without damage. Next, the same thresholds were applied to flight data from an OH-58 helicopter transmission. Results showed that thresholds defined in test rigs can be used to define thresholds in flight to correctly classify the transmission operation as normal.

Introduction

The goal in the development of diagnostic tools used for fault detection of helicopter transmissions is to provide real-time performance monitoring of aircraft operating parameters and to be highly reliable to minimize false alarms. Various diagnostic tools exist for diagnosing damage in helicopter transmissions, the most common being vibration-based tools. Using vibration data collected from gearbox accelerometers, algorithms are developed to detect when gear damage has occurred.

Over the past 25 years, numerous vibration-based algorithms for gear damage detection have been developed. In order to evaluate the performance of individual vibration based diagnostic tools, a set of standard thresholds must be defined. When defining the threshold or limit of a diagnostic tool there is a tradeoff between the sensitivity of the limit to indicate damage and the number of false alarms. If a limit is decreased, damage may be detected, but more false alarms may result. If a limit is increased, false alarms may decrease, but the algorithms will be less sensitive to damage.

The simplest approach of setting thresholds for vibration diagnostic tools is to gather baseline data under "normal" operating conditions, and set the threshold to values that exceed "normal" operating conditions. HUMS (Health Usage Monitoring Systems) manufacturers' analysis of fleet data have observed significant variances of indicator levels between gearbox components [1]. Due to limited damage data in flight, diagnostic tools must be developed in controlled ground test environments. Defining thresholds for different types of rig component failures is required for predicting future helicopter component failures. The objective of this research is to assess the performance of vibration based diagnostic tools on both flight data and test rig data using thresholds defined in a test rig environment. The flight data were collected from an OH-58C helicopter. The test rig data were collected from the NASA Glenn Spur Gear Fatigue Rig. The threshold assessment will be applied to experimental data collected in NASA Glenn test rigs under normal conditions and damage progression conditions and to data collected under normal conditions on the helicopter. Relating the performance of rig data to flight data with standard thresholds will provide valuable information on relating the operational effects of flight to test rig data. This information can then be used to improve the performance of the diagnostic tool by identifying damage detection thresholds in test rigs that have low false alarm rates on helicopters.

Presented at the American Helicopter Society 59th Annual Forum, Phoenix, Arizona, May 6–8, 2003.

Experimental Set-up and Procedures

All fatigue tests were conducted using the Spur Gear Fatigue Test Rig facility located at NASA Glenn Research Center (GRC). The spur gear test rig is capable of running gears, under high speeds and loads, until pitting damage is detected [2]. Figure 1 shows the test apparatus in the facility and a photo of the gearbox with the cover removed. Operating on a four square principle, the shafts are coupled together with torque applied by a hydraulic loading mechanism that twists two shafts with respect to one another. The power required to drive the system was just high enough to overcome friction losses in the system [3]. The test gears used were standard spur gears with 28 teeth, 8.89 cm pitch diameter, and 0.64 cm face width. The type of damage under investigation during fatigue tests is pitting damage. An example of this type of damage on a gear tooth is shown on Figure 1. Pitting is a fatigue failure that occurs when small pieces of material break off from the gear surface, producing pits on the contacting surfaces [4]. Gears are run until pitting occurs on one or more teeth.

Data were collected using vibration, oil debris, speed and pressure sensors installed on the test rig. Vibration was measured through the gearbox shaft using a miniature, lightweight, piezoelectric accelerometer. Location of this accelerometer is shown on Figure 1. Oil debris data were collected using a commercially available oil debris sensor that measures the change in a magnetic field caused by passage of a metal particle where the amplitude of the sensor output signal is proportional to the particle mass [5]. Shaft speed was measured with an optical sensor that creates a pulse signal for each revolution of the shaft. Load pressure was measured using a capacitance pressure transducer. Data were collected once per minute and processed by a data acquisition system program named ALBERT, Ames-Lewis Basic Experimentation in Real Time, co-developed by NASA Glenn and NASA Ames. The time-synchronous averages of the vibration data were calculated. Synchronous averaging of time signals is a technique used to extract periodic waveforms from additive noise by averaging the vibration signal over one revolution of the shaft. The signal time-synchronous average is obtained by taking the average of the signal in the time domain with each record starting at the same point in the cycle as determined by the once per revolution tachometer signal.

Flight data were collected at NASA Ames Research Center (ARC) from an OH-58 helicopter transmission. Vibration data were collected from 6 accelerometers, three uni-axial and one tri-axial, mounted on the transmission housing. Shaft speed was measured by a once per revolution sensor. The data were collected from 14 maneuvers performed by an OH-58C Kiowa helicopter [6, 7]. A description of the 14 maneuvers is listed in Table 1. Data were collected for 34 seconds for 12 repetitions of each maneuver. The vibration diagnostic algorithms for this analysis focused on the health of the 19-tooth pinion on the input shaft of the main rotor transmission. Figure 2 shows the accelerometer locations. The time-synchronous averages of the vibration data were calculated from the flight data. For each 34-second maneuver, 48 time synchronous averages were calculated.

For both the rig data and the flight data, two parameters, FM4 and M8A, were calculated from the time synchronous averaged vibration data. Table 2 lists the equations used to calculate FM4 and M8A [8, 9]. Figure 3 is a representative example of the data that shows FM4 measured by accelerometers 1 through 6 for maneuver A. The segmented lines indicate the 12 repetitions of Maneuver A. Each repetition consists of the 48 time synchronous averages for each data set.

Analysis Discussion and Results

The assessment process consists of defining thresholds that indicate spur gear damage during rig tests and also minimize false alarms when applied to rig and helicopter gears with no damage. Three sets of data were evaluated. One set contained 5 experiments with damage on spur gears. One set contained 3 experiments with no damage to spur gears. One set contained the OH58 flight data collected during the 14 maneuvers listed in Table 1.

FM4 and M8A maximum values during spur rig experiments with damage are shown in Tables 3 and 4. During tests, the rig was shut down at inspection intervals, and damage progression was documented with a video inspection system that consists of a micro camera inserted in the gearbox viewing ports. The table indicates the reading number when video inspection was performed. The highlighted cells indicate when pitting was first observed.

Thresholds for FM4 were defined using this data, and 18 additional spur rig experiments, in a previous research effort for input into a data fusion model [10]. FM4 and M8A time history data (reading = 1 minute) plotted for experiments 1 through 5 are shown in Figures 4 and 5. The plots are separated into 3 sections based on inspection intervals. The first interval indicated by green in when no damage was observed. The second interval indicated by yellow is when damage occurred. The third interval, indicated by red, is when damage was observed via inspection. For example, referring to Table 3, experiment 1, the

inspection interval where damage occurred begins at reading 1574 and ends at reading 2199 when damage was first observed. Data analysis of the thresholds will focus on both false alarm rates in the intervals prior to damage being observed, and the ability of the algorithm to detect damage during the interval damage was observed.

The maximum, minimum, mean and standard deviation values for FM4 and M8A spur rig experiments with no damage are shown in Tables 5 and 6. At test completion, no damage was observed on the gear teeth. FM4 and M8A data plotted for experiments 6 through 8 are shown in Figures 6 and 7. Data analysis of the thresholds will focus on false alarms indicated for these three experiments.

The maximum, minimum, mean and standard deviation values for FM4 and M8A during 12 repetitions of OH58 flight maneuvers A through L for accelerometer a6 is shown in Tables 7 and 8. All data were collected on a healthy helicopter transmission. Performance evaluation of the flight data will be limited to minimizing false alarms to correctly classify the transmission operation as normal.

Evaluating diagnostic tool performance of FM4 and M8A depends on the thresholds defined to indicate levels of damage. Defining threshold limits for vibration algorithms to indicate when damage occurs is a challenging task. Although different thresholds are identified in the literature, they are specific to the test environment. Thresholds levels are selected so that FM4 and M8A values are below the threshold for gears in good condition and above the threshold for damaged gears [8, 11–16]. In some published cases, the magnitude of FM4 did not increase significantly for all experiments when damage occurred, falling below the chosen threshold [12, 14].

One rule of thumb for setting limits on performance parameters is to set the limit to 3 times the standard deviation of the mean [17]. Defining an amplitude limit based on the mean plus 3 standard deviations has also been used for helicopter drive train vibration monitoring systems [18]. Reviewing the rig and flight mean and standard deviation data for FM4 and M8A, listed in Tables 5 through 8, using the mean plus 3 standard deviations would result in numerous false alarms, and thus is a poor choice for a threshold.

Using fuzzy logic membership functions to define limits showed promising results for minimizing false alarms for spur gears and spiral bevel gears [19]. FM4 membership values were defined by looking at the maximum FM4 value within each inspection interval. Membership values are defined for 2 levels of damage: damage low (DL) and damage high (DH). The membership function for FM4 is shown in Figure 8. Limits identified during this analysis were used to evaluate the performance of FM4 and M8A under different operating environments

The false alarm rates for 5 thresholds during spur rig experiments for FM4 and M8A are listed in Tables 9 and 10. For the experiments with damage, the false alarm rates were defined for the inspection intervals where damage did not occur. For example, referring to Table 3, Experiment 1, the false alarm rates are only calculated for readings 1 through 1573. The false alarm rates are the percentage of time a threshold is exceeded during an experiment. The number of readings the false alarms are based on are listed in the bottom row of both tables. Readings for the rig experiment were taken every minute. It is very clear that setting the threshold higher eliminates false alarms. However, if the limit is set too high to minimize false alarms, damage may never be detected. If the limit is not set sensitive enough to detect damage, the diagnostic tool is essentially worthless. In addition, unless progressive damage data is available, setting these limits becomes a challenging task.

The usefulness of a threshold for damage detections depends upon detecting the damage as well as avoiding false alarms. Tables 11 and 12 show the damage detection sensitivity from the GRC test rig data. The reading numbers listed in the tables indicate the difference between the reading when damage was detected and the reading when damage was observed. The readings in the last row of the table indicate the first reading number after the gear was last observed to contain no damage (see Tables 3 and 4). Although the gears are inspected periodically for damage, no on-line inspection system currently exists that can determine the exact instance that damage occurred. For the spur rig tests, the on-line oil debris sensor was used to indicate shutdowns based on the amount of debris measured in the lubrication line. Unlike the false alarm tables, for Tables 11 and 12, larger values indicate poor performance. Large values indicate the amount of time to detect damage increased. For example, reviewing Table 11 for experiment 3, damage occurred between reading 519 and 2065. If the threshold for FM4 was set at 5.18, the metric did not increase in value for 2575 readings or 43 hours. The “N/A” in Tables 11 and 12 indicates thresholds were never exceeded, although experiment 4 ended when damage was first observed. For minimum false alarms FM4 should be set to 4.04 and M8A set at 394 in the test rig.

Threshold values for FM4 and M8A were defined based on test rig data. The next step is to determine if these

lab-based thresholds can be used in the aircraft. Reviewing Tables 13 and 14, FM4 set to 4.04 and M8A set to 394 results in minimal false alarms for all 14 maneuvers. Since no damage occurred to the gearbox during flight experiments, damage detection intervals could not be calculated for this experiment. Therefore, until pitting damage occurs, and FM4 and M8A data is collected when damage occurs, this sensitivity of FM4 and M8A to damage cannot be verified.

After reviewing this data, it becomes very clear that keeping the threshold high will indicate excellent performance of a vibration based metric as long as no damage occurs. However, damage will never be detected if the limit is high. Although mean and standard deviation data is often used to define limits for vibration-based metrics, reviewing Figures 4 and 5, it is clear that their usefulness is limited. In some cases the algorithm shows very little change when damage first appears and decreases in value when damage progresses.

One statistical approach was reviewed to determine if the signature data from the rig resembled the data from the aircraft during normal conditions. For this approach, test rig data were analyzed by plotting the estimated probabilities of the relative frequency plots (PDF) on Figures 9 and 10, showing FM4 and M8A respectively for the GRC test rig experiments. For each comparison plot, histograms were formed using a heuristically determined number of bins between the minimum and maximum metric values. The estimated PDF is equal to the histogram divided by the number of samples multiplied by the histogram bin width:

$$PDF = \frac{histogram}{(readings) \times (bin\_width)}$$

The histogram bin width is equal to the maximum value of FM4 or M8A minus the minimum value divided by the number of bins.

$$bin\_width = \frac{FM4_{\max} - FM4_{\min}}{\#bins}$$

The number of bins is based on the number of readings in the experiment for each state, no damage, unknown, damage. When more data was available, more bins were used with the goal of providing enough bins to see the shape of the distribution, but not so many as to produce a jagged curve. The histograms were rescaled so that the integral under the estimated distribution is 1 as described above. The semi-log scale more clearly shows differences in the high-end tails for cases with damaged and undamaged gears.

The 3 states (no damage, unknown, damage) were plotted using 3 different colors for the five experiments with damage occurring on the gear. The unknown state is the inspection interval when damage occurred, but when the exact reading damage occurred was not verified with video inspection. The sixth plot shows the 3 experiments without damage and the no damage data from the other 5 plots. In the cases where no damage occurred, the distribution curves display similar shape with a peak near or below 3 for FM4 and near or below 100 for M8A. The high-end tails of the distributions for the cases with no damage drop off fairly smoothly. In the cases where gear damage is known to be present in experiments 1, 3, and 5, the shape of the distribution differs from the cases without gear damage; the upper end of the distribution curves do not drop off as fast and sometimes contain relatively flat spots or increases. The curves for experiment 2 display less difference with and without damage than in the other experiments. There is no data for experiment 4 with known damage.

Figure 11 displays the estimated probability distribution curves for the ARC flight measurements of FM4 and M8A respectively. Note that the shapes and levels are quite similar to the curves for the GRC test rig data without gear damage. In all cases, the high levels of the metrics relative to suggested thresholds occur with low frequency suggesting a low probability.

If a threshold is selected to ensure detection of damage based upon the test rig data, some flight measurements and some measurements from test rig data without damage exceed that threshold. A better damage detection scheme might be created with the metrics by testing for the different shapes and locations of the estimated probability distribution functions with and without gear damage. Two types of tests come to mind. First, metric data could be tested with the Kolmogorov-Smirnoff test for comparisons of the cumulative distribution function. The other is that a neural net could be trained to distinguish between damaged and undamaged gears with inputs consisting of selected inverse values for the cumulative distribution function. More work is needed to explore the value of the two possible tests for damage and to determine specific testing issues such as the size of the sample used to estimate the distribution function.

Conclusions

A method to set fault threshold limits for FM4 and M8A in a test rig was used to set thresholds in a helicopter transmission. Setting standard thresholds for vibration-based algorithms provides a method to weigh the diagnostic tool performance based on their individual strengths and weaknesses prior to integration into a propulsion health management system. The thresholds

can be optimized for minimum false alarms or maximum sensitivity to damage. No thresholds for FM4 or M8A were found to give both high sensitivity to damage and no false alarms in this study. Additional flight data is required to verify damage detection sensitivity can be maintained in a flight environment. A database of failure progression data for other types of failures would enable application of this process to failure mechanisms in addition to pitting fatigue failure.

The NASA Glenn Spur Gear Fatigue Test Rig was used to collect vibration data on spur gears with and without pitting damage. Vibration data were also collected from an OH-58 helicopter transmission in flight. Thresholds for vibration-based gear damage detection algorithms, FM4 and M8A were analyzed on the rig and the flight data. Based on this analysis, the following conclusions can be made:

1. In the spur gear fatigue test rigs, threshold values of 4.04 for FM4 and 394 for M8A result in the minimum false alarms, while maintaining sensitivity to gear pitting when damage occurred.
2. False alarms only occurred during maneuver B when a threshold of 4.04 was set for FM4 during flight maneuvers. False alarms occurred during maneuvers B and F if M8A is set at 394. However, damage data was not available to identify their sensitivity to damage.
3. The histograms indicate that the probability distribution curves from flight data are very similar to the distributions for rig data under no damage conditions.
4. Distributions of the data change when damage occurs indicating testing for distributions may be an alternative to setting thresholds.

References

[1] Larder, B.D.: An Analysis of HUMS Vibration Diagnostic Capabilities. American Helicopter Society 53rd Annual Forum, May 1997.

[2] Scibbe, H.W., Townsend, D. P., and Aron, P.R: Effect of Lubricant Extreme Pressure Additives on Surface Fatigue Life of AISI 9310 Spur Gears. NASA Technical Paper 2408, December 1984.

[3] Lynwander, Peter: Gear Drive Systems: Design and Application. Marcel Dekker, New York, NY, 1983.

[4] Townsend, Dennis P.: Dudley's Gear Handbook. McGraw-Hill, New York, NY, 1991.

[5] Howe, B.; and Muir, D.: In-Line Oil Debris Monitor (ODM) for Helicopter Gearbox Condition Assessment. AD–a347 503. Defense Technical Information Center, Ft. Belvoir, VA, 1998.

[6] Mosher, M., Pryor, A.H., and Huff, E.M.: Evaluation of Standard Gear Metrics in Helicopter Flight Operation, 56th Meeting of the MFPT, April 2002.

[7] Huff, E. M.; Tumer, I. Y.; and Mosher, M.: An Experimental comparison of Transmission Vibration Responses from HO-58 and AH-1 Helicopters. American Helicopter Society 57th Annual Forum, May 2001.

[8] Stewart, R.M.: Some Useful Data Analysis Techniques for Gearbox Diagnostics. Machine Health Monitoring Group Report MHM/R/10/77, University of Southampton, 1977.

[9] Martin, H.R. Statistical Moment Analysis as a Means of Surface Damage Detection. Proceedings of the 7th International Modal Analysis Conference, Society for Experimental Mechanics, Schenectady, NY, 1989, pp. 1016-1021.

[10] Dempsey, Paula J.: Integrating Oil Debris and Vibration Measurements for Intelligent Machine Health Monitoring. Ph.D. Thesis, The University of Toledo, May 2002.

[11] Forrester, B.D.: Time Frequency Domain Analysis of Helicopter Transmission Vibration. Department of Defense DSTO AR-005-606, August 1991.

[12] Zakrajsek, J.J.; Townsend, D.P.; and Decker, H.J.: An Analysis of Gear Fault Detection Methods as Applied to Pitting Fatigue Failure Data. NASA TM–105950, January 1993.

[13] Choy, F.K.; Huang, S.; Zakrajsek, J.J; Handschuh, R.F.; and Townsend, D.P.: Vibration Signature Analysis of a Faulted Gear Transmission System. NASA TM-106623, June 1994.

[14] Zakrajsek, James J.; Decker, H.J.; Handschuh, R.F.; and Lewicki, D.G.: Detecting Gear Tooth Fracture in a High Contact Ratio Face Gear Mesh. NASA TM–106822, January 1995.

[15] Dempsey, Paula J.: A Comparison of Vibration and Oil Debris Gear Damage Detection Methods Applied to Pitting Damage. NASA/TM—2000-210371, September 2000.

[16] Campbell, R.L.; Byington, C.S.; and Lebold, M.S.: Generation of HUMS Diagnostic Estimates Using Transitional Data. Proceedings of 13th International Congress on Condition Monitoring and Diagnostic Engineering Management, Henry C. Pusey and Raj B.K.N. Rao, eds., The Society for Machinery Failure Prevention Technology, Haymarket, VA, pp. 587-595, April 2000.

[17] Slemp, M., Skeirik, R.: Using Statistics to Avoid Vibration False Alarms. P/PM Technology, February 1999.

[18] Dousis, D.A., Giles, L.B., Hale, R.K. and Priest, T.B: V-22 Drive Train Vibration Monitoring System Ground Station Development. American Helicopter Society 55th Annual Forum, May 1999.

[19] Dempsey, P.J. and Afjeh, A.A.: Integrating Oil Debris and Vibration Gear Damage Detection Technologies Using Fuzzy Logic. American Helicopter Society International Forum 58, June 2002.

Table 1. OH58 Flight Maneuvers

Maneuver		Maneuver	
A	Level, forward ~5% torque	H	Hover, ~10 ft
B	Level, forward ~80% torque	I	Hover~10 ft, turn left
C	Level, sideways left	J	Hover~10 ft, turn right
D	Level sideways right	K	20… bank left turn
E	Climb, ~55% torque	L	20… bank right turn
F	Descent, ~10% torque	M	Climb, ~80% torque
G	Flat pitch on ground	N	Descent, 35% torque

Table 2. Formulas for FM4 and M8A

Metric	Filtering	Formula	Numerator	Denominator	Nominal Value
FM4	Difference, remove gmhs, first order side bands, 1/rev, 2/rev	$\frac{\frac{1}{N}\sum_{n=1}^{N}\left(d_n-\bar{d}\right)^4}{\left[\frac{1}{N}\sum_{n=1}^{N}(d_n-\bar{d})^2\right]^2}$	Fourth moment about mean of difference signal	Square of variance of difference signal	3
M8A	Difference, remove gmhs, first order sidebands, 1/rev, 2/rev	$\frac{\frac{1}{N}\sum_{n=1}^{N}\left(d_n-\bar{d}\right)^8}{\left[\frac{1}{N}\sum_{n=1}^{N}(d_n-\bar{d})^2\right]^4}$	Eighth moment about mean of difference signal	Fourth power of variance of difference signal	105

Table 3. Spur Rig FM4 Max at Video Gear Inspection Intervals

Experiment 1		Experiment 2		Experiment 3		Experiment 4		Experiment 5	
Rdg#	FM4 Max	Rdg#	FM4 Max	Rdg#	FM4 Max	Rdg#	FM4 Max	Rdg#	FM4 Max
1573	3.49	58	3.51	64	3.17	62	3.17	60	3.91
2199	5.23	2669	4.66	150	3.04	1405	3.28	2810	3.78
2296	5.03	2857	5.91	378	3.97	2566	3.10	2885	3.35
2444	5.84	3029	3.86	518	2.94	4425	3.68	2957	3.29
				2065	2.86			9328	3.50
				2366	4.19			12061	3.66
				3671	2.90			12368	4.13
				4655	5.43				
				4863	5.49				

Note: Highlighted cells identify reading when destructive pitting was first observed

Table 4. Spur Rig M8A Max at Video Gear Inspection Intervals

Experiment 1		Experiment 2		Experiment 3		Experiment 4		Experiment 5	
Rdg#	M8A Max	Rdg#	M8A Max	Rdg#	M8A Max	Rdg#	M8A Max	Rdg#	M8A Max
1573	261.77	58	262.92	64	134.49	62	205.40	60	378.36
2199	929.78	2669	607.07	150	68.36	1405	132.72	2810	319.66
2296	547.31	2857	896.84	378	408.27	2566	164.90	2885	166.94
2444	727.05	3029	345.12	518	99.25	4425	394.88	2957	137.16
				2065	98.25			9328	147.57
				2366	417.07			12061	226.06
				3671	88.93			12368	440.04
				4655	564.45				
				4863	560.70				

Note: Highlighted cells identify reading when destructive pitting was first observed

Table 5. Spur Rig FM4 Max–No Gear Damage

Experiment	Readings	FM4 Min	FM4 Max	FM4 Mean	FM4 Std Dev
1	1573	2.13	3.49	2.71	0.1778
2	58	2.64	3.51	2.94	0.1975
3	518	1.98	3.97	2.52	0.2155
4	2566	2.27	3.28	2.84	0.1629
5	9328	2.28	3.91	3.05	0.1644
6	10000	2.35	4.61	3.03	0.2486
7	8000	2.32	4.18	3.07	0.2852
8	21066	2.09	5.17	2.98	0.3487

Table 6. Spur Rig M8A Max–No Gear Damage

Experiment	Readings	M8A Min	M8A Max	M8A Mean	M8A Std Dev
1	1573	26.46	261.77	73.22	26.06
2	58	35.16	262.92	83.64	43.92
3	518	16.14	408.27	47.44	23.88
4	2566	32.01	205.40	82.71	23.94
5	9328	31.51	378.36	63.55	17.32
6	10000	29.67	716.24	117.87	67.62
7	8000	26.68	318.97	74.68	29.71
8	21066	16.03	325.81	51.67	20.87

Table 7. OH58 Flight FM4 Maneuvers–A6

Maneuver	Readings	FM4 Min	FM4 Max	FM4 Mean	FM4 Std Dev
A	576	1.88	3.43	2.60	0.3779
B	576	2.27	4.92	2.91	0.3791
C	576	1.91	3.89	2.83	0.3629
D	576	2.01	3.44	2.68	0.2155
E	576	1.90	3.30	2.59	0.3321
F	576	2.18	3.79	2.73	0.2573
G	768	2.41	3.55	2.89	0.2135
H	768	2.19	3.68	2.71	0.2708
I	576	2.26	3.60	2.77	0.2530
J	576	2.28	3.77	2.75	0.2550
K	576	2.19	3.46	2.86	0.1850
L	576	2.32	3.60	2.88	0.1974
M	576	2.23	3.76	2.67	0.2267
N	576	1.87	2.70	2.14	0.1371

Table 8. OH58 Flight M8A Maneuvers–A6

Maneuver	Readings	M8A Min	M8A Max	M8A Mean	M8A Std Dev
A	576	12.59	165.83	55.05	26.89
B	576	24.47	1212.82	101.52	92.98
C	576	13.48	238.79	76.21	43.83
D	576	17.67	190.33	57.35	22.56
E	576	12.69	143.32	52.28	24.83
F	576	22.97	461.34	70.41	44.00
G	768	35.04	153.92	72.42	19.37
H	768	25.41	237.27	61.03	28.87
I	576	24.72	231.57	65.23	28.68
J	576	25.89	278.87	63.99	31.40
K	576	25.60	201.38	72.93	23.16
L	576	30.21	202.67	76.66	27.57
M	576	23.57	214.12	54.01	22.85
N	576	11.96	62.43	24.12	8.73

Table 9. Spur Rig False Alarm Rates (%)

FM4	Experiments							
Thresholds	1	2	3	4	5	6	7	8
5.18	0	0	0	0	0	0	0	0
4.04	0	0	0	0	0	0.12	0.06	1.03
3.66	0	0	0.19	0	0.13	1.70	1.90	4.43
3.07	2.35	24.14	0.77	7.52	43.76	39.78	50.08	34.35
Readings	1573	58	518	2566	9328	10000	8000	21066

Table 10. Spur Rig False Alarm Rates (%)

M8A	Experiments							
Thresholds	1	2	3	4	5	6	7	8
716	0	0	0	0	0	0.01	0	0
394	0	0	0.19	0	0	0.76	0	0
350	0	0	0.19	0	0.02	1.37	0	0
300	0	0	0.19	0	0.04	2.77	0.03	0.01
226	0.06	1.72	0.19	0	0.11	6.70	0.30	0.04
Readings	1573	58	518	2566	9328	10000	8000	21066

Table 11. Damage Detection Interval
(Reading damage detected–Reading damage observed)

FM4	Experiments				
Thresholds	1	2	3	4	5
5.18	0	74	2575	N/A	N/A
4.04	0	0	16	N/A	56
3.66	0	0	16	0	0
3.07	0	0	1	0	0
Initial Reading	1574	59	519	2567	9329

Table 12. Damage Detection Interval
(Reading damage detected–Reading damage observed)

M8A	Experiments				
Thresholds	1	2	3	4	5
716	0	2684	N/A	N/A	N/A
394	0	0	1562	1836	2789
350	0	0	1562	1836	2789
300	0	0	1562	1836	2784
226	0	0	1562	1836	495
Initial Reading	1574	59	519	2567	9329

Table 13. OH58 Flight False Alarm Rates (%)

FM4 Thresholds	Maneuvers													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
5.18	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4.04	0	0.52	0	0	0	0	0	0	0	0	0	0	0	0
3.66	0	4.51	1.04	0	0	0.87	0	0.39	0	0.35	0	0	0.17	0
3.07	4.69	24.43	24.13	24.13	4.17	1.56	20.31	10.03	12.85	11.46	13.37	17.71	6.60	0
Readings	576	576	576	576	576	576	768	768	576	576	576	576	576	576

Table 14. OH58 Flight False Alarm Rates (%)

M8A Thresholds	Maneuvers													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
716	0	0.35	0	0	0	0	0	0	0	0	0	0	0	0
394	0	0.87	0	0	0	0.35	0	0	0	0	0	0	0	0
350	0	1.56	0	0	0	0.52	0	0	0	0	0	0	0	0
300	0	3.13	0	0	0	0.69	0	0	0	0	0	0	0	0
226	0	7.99	0.35	0	0	1.56	0	0.26	0.17	0.69	0	0	0	0.17
Readings	576	576	576	576	576	576	768	768	576	576	576	576	576	576

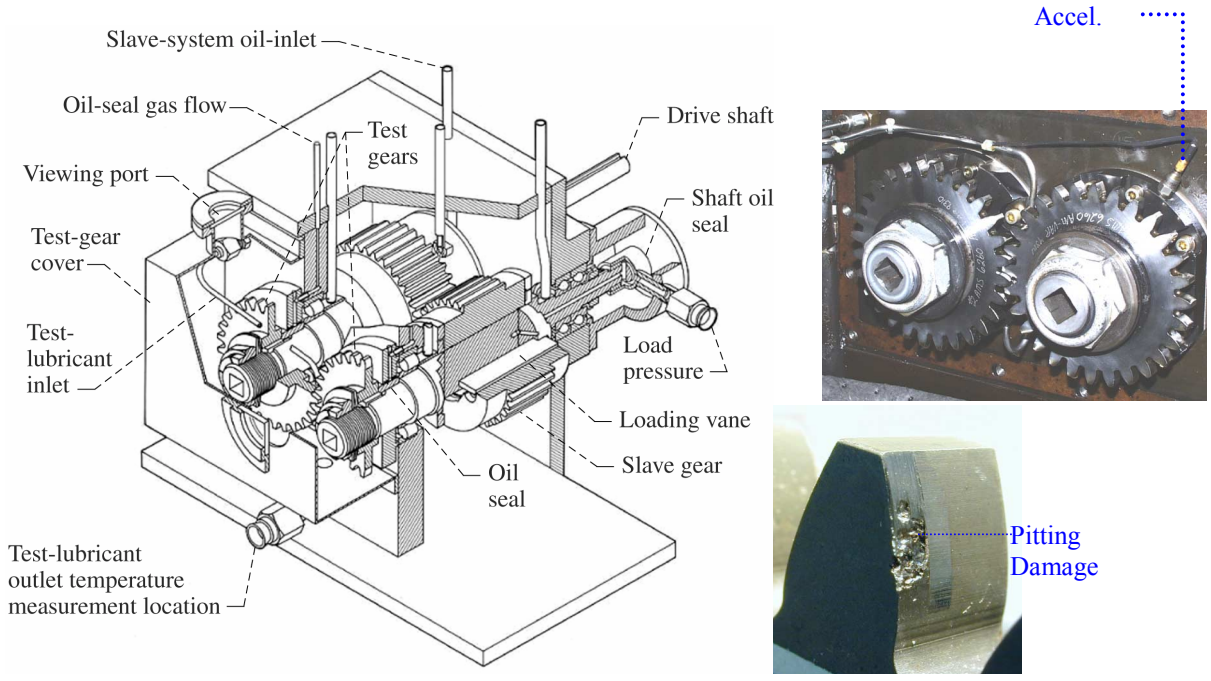

Figure 1. Spur Gear Fatigue Test Rig

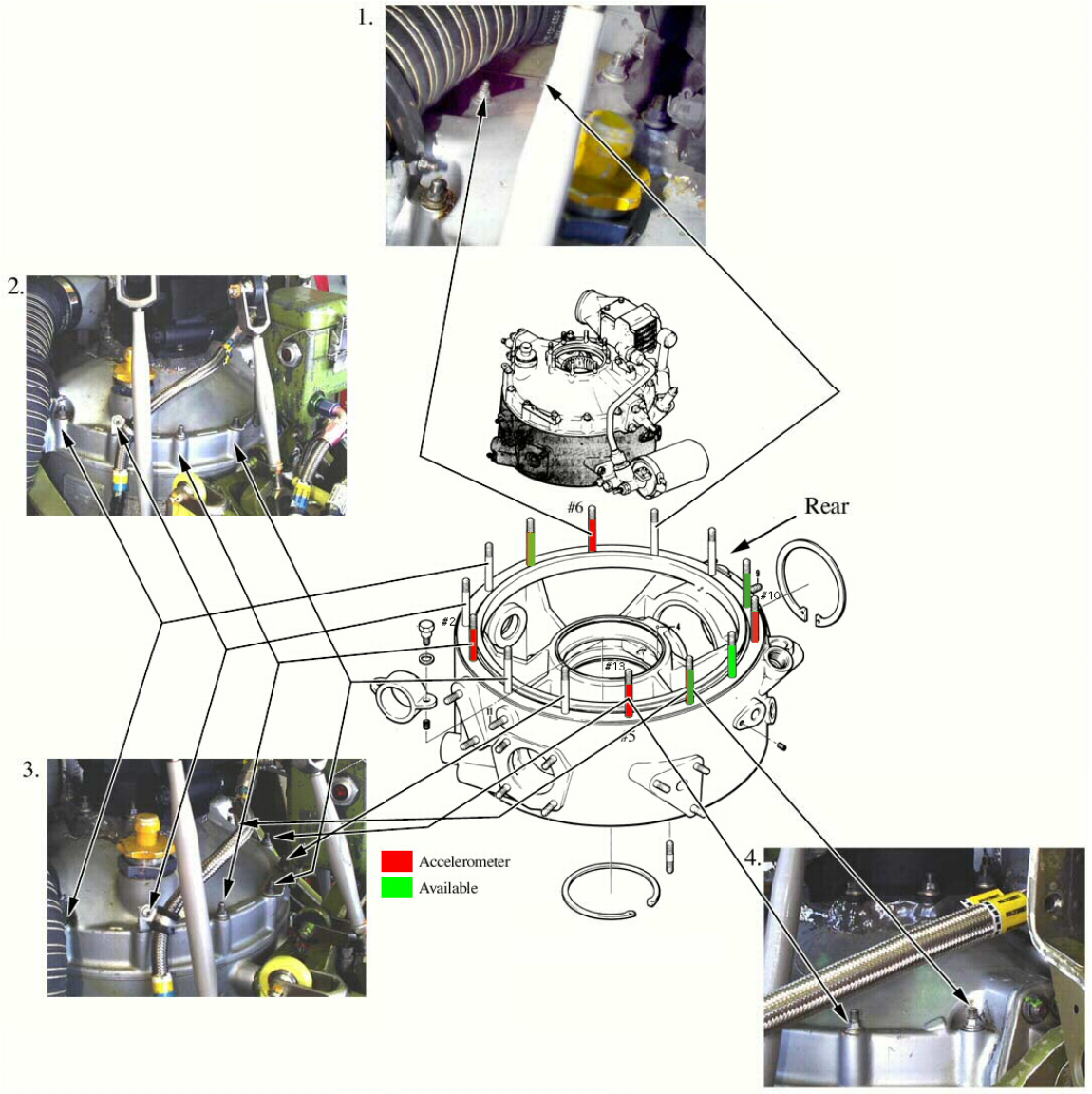

Figure 2. Accelerometer Locations on OH58

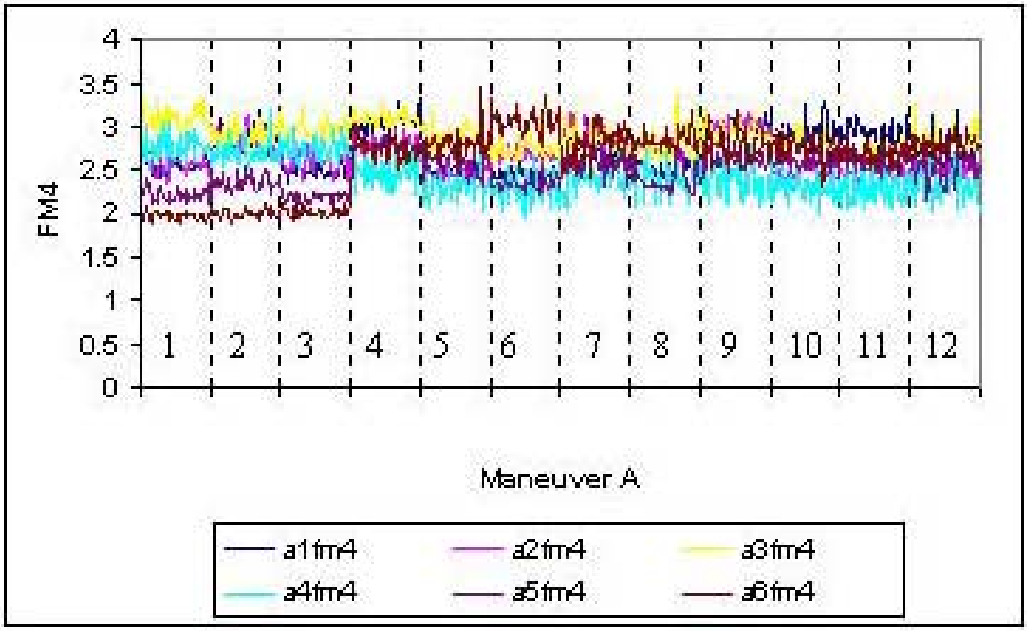

Figure 3: FM4 for 12 Repetitions of Maneuver A

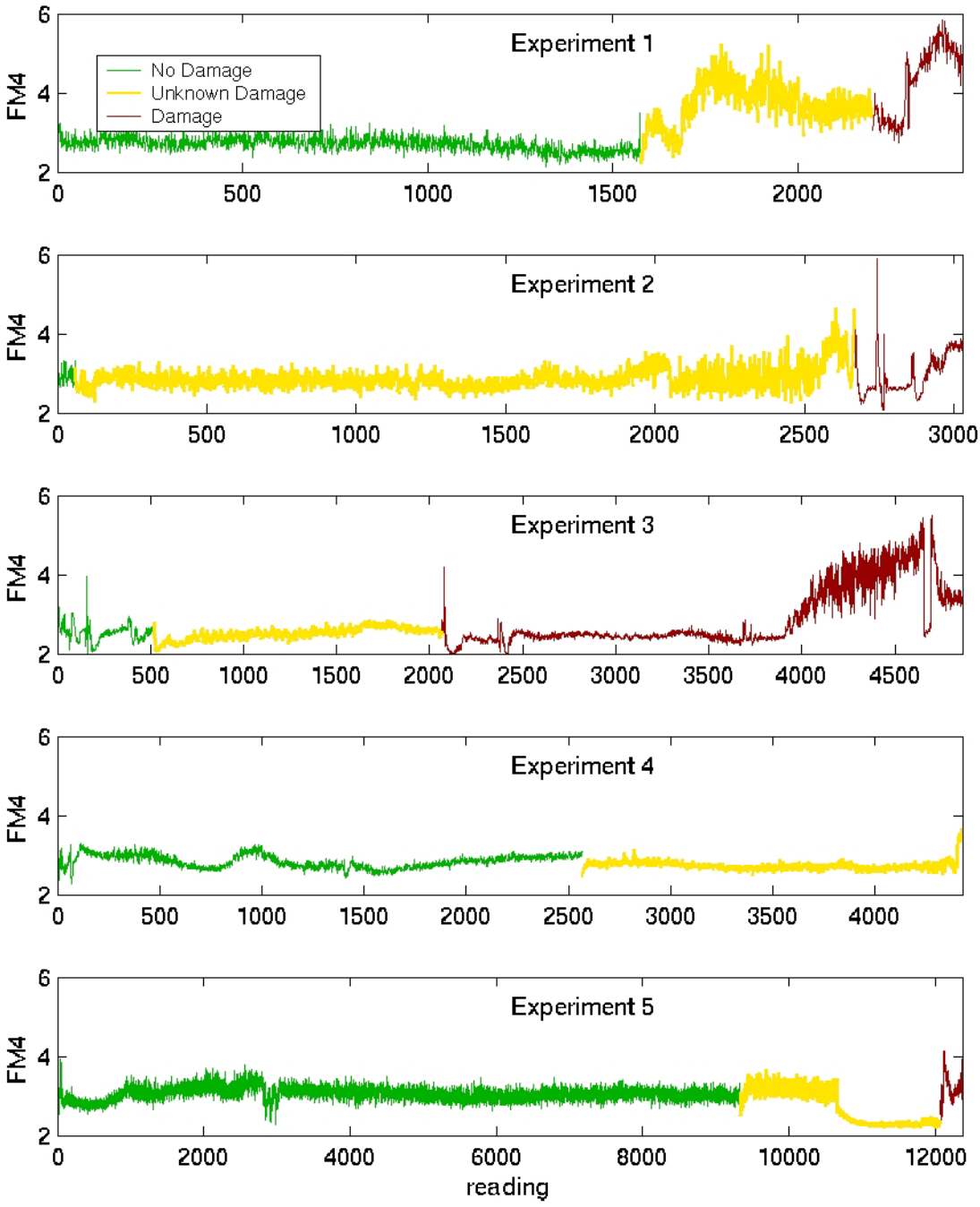

Figure 4. FM4-Spur Rig Experiments 1 through 5

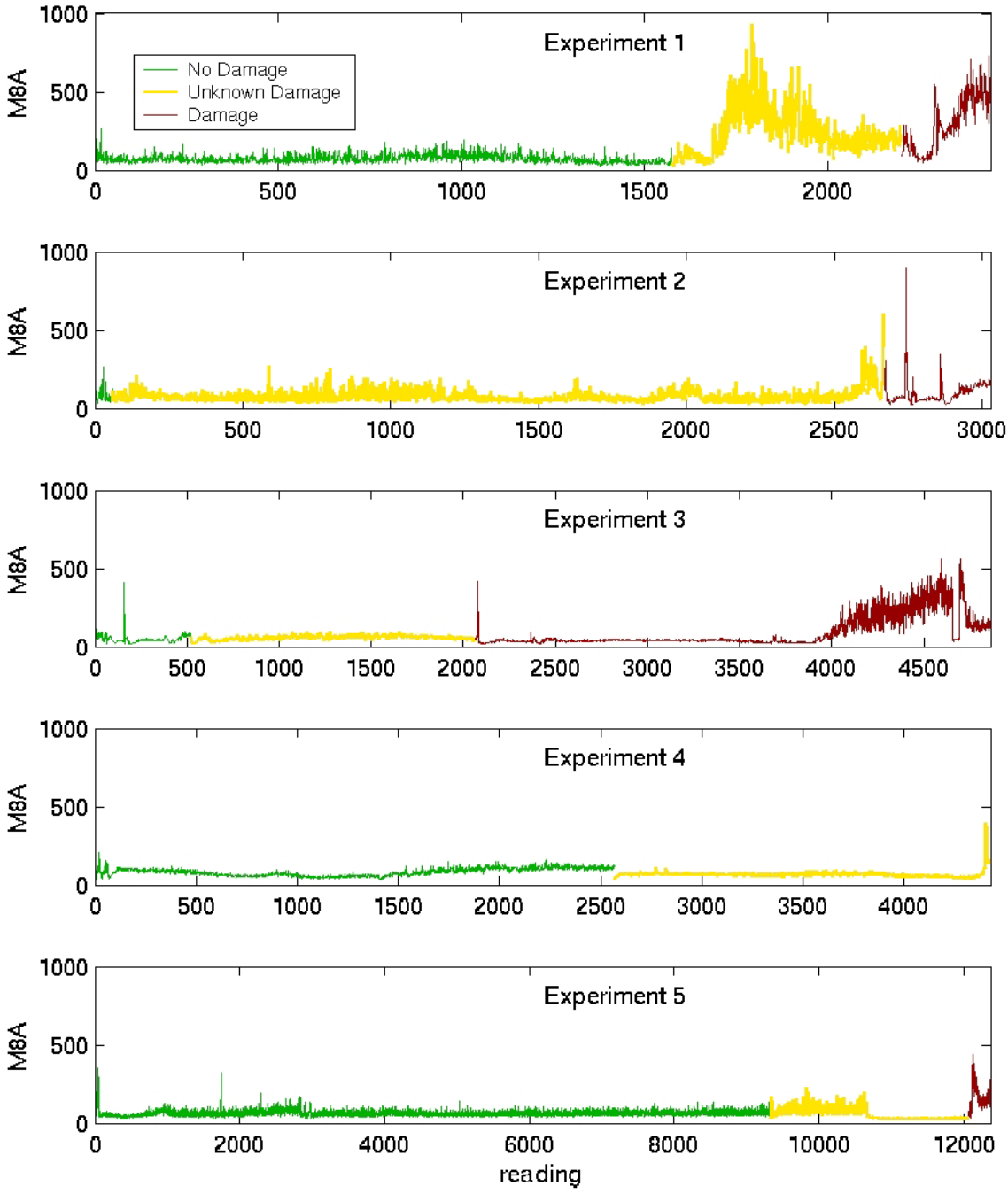

Figure 5. M8A-Spur Rig Experiments 1 through 5

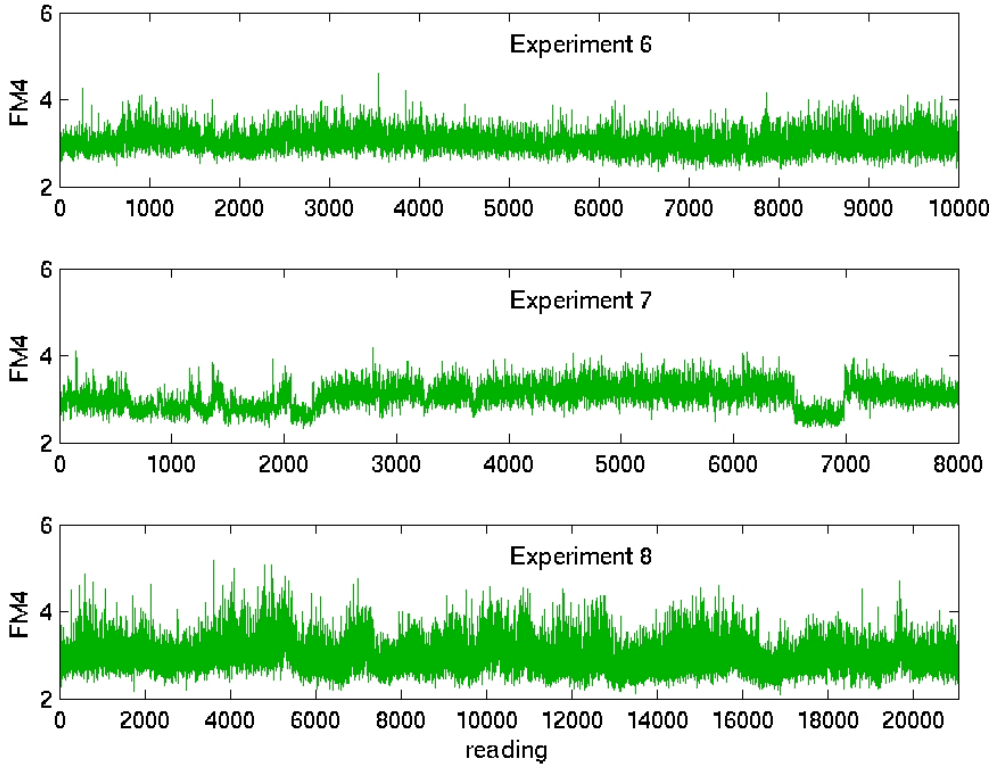

Figure 6. FM4-Spur Rig Experiments 6 through 8

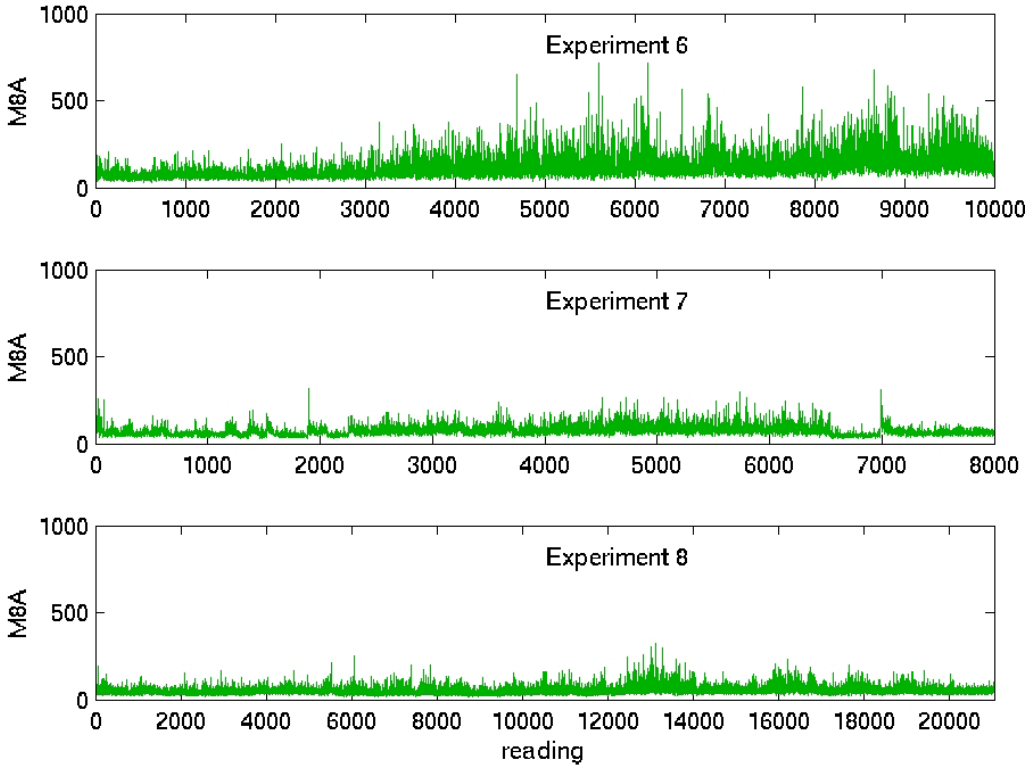

Figure 7. M8A-Spur Rig Experiments 6 through 8

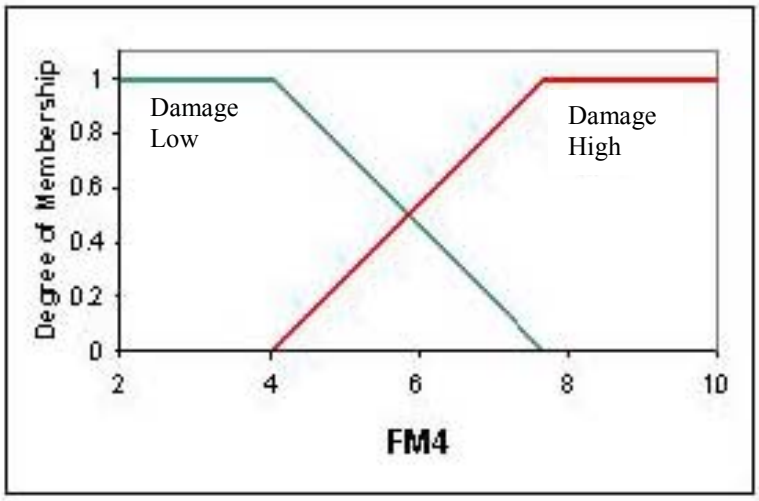

Figure 8. Membership Function for FM4 [10]

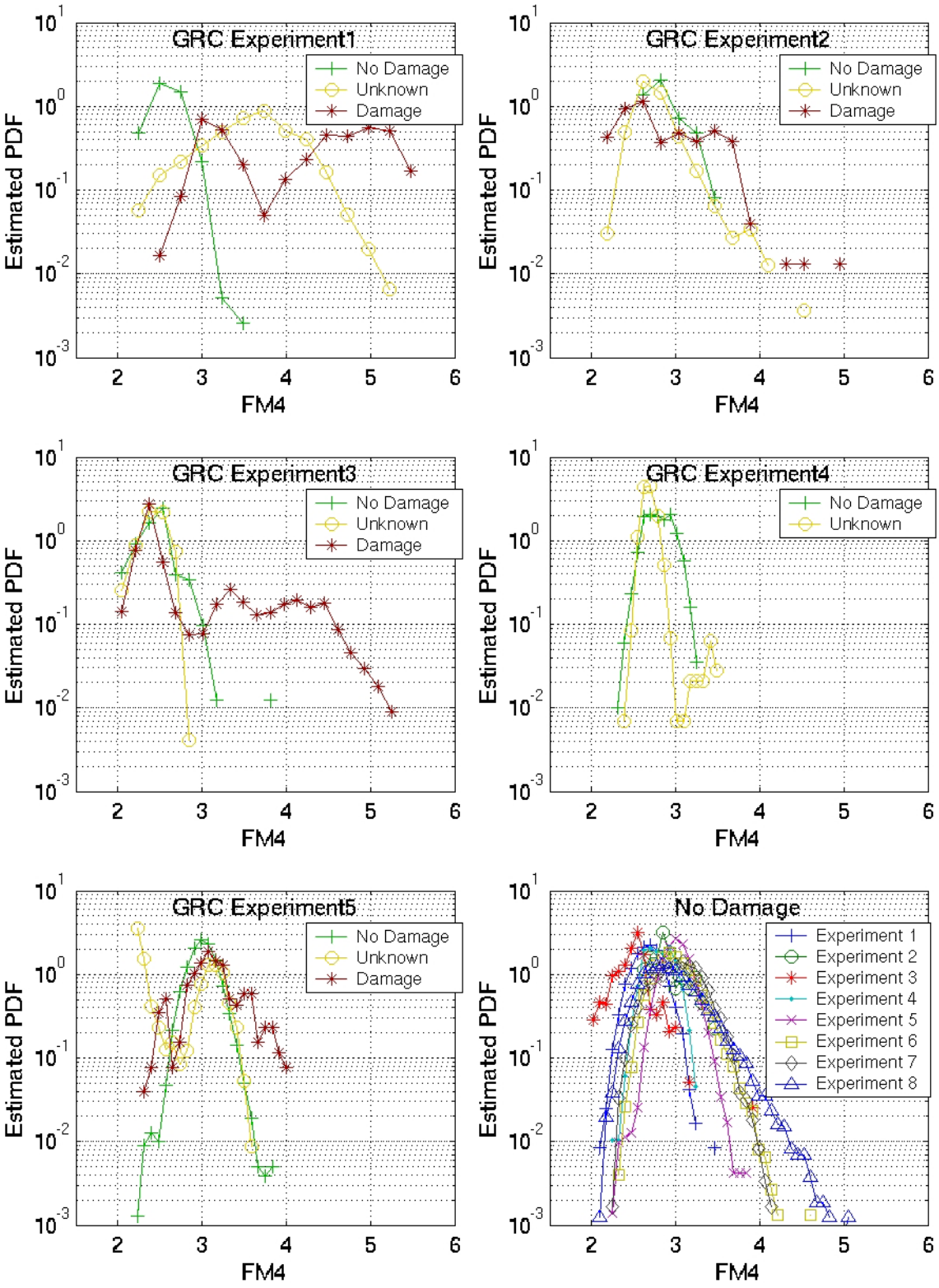

Figure 9. Estimated Probability Distribution Functions from GRC Test Rig Experiments

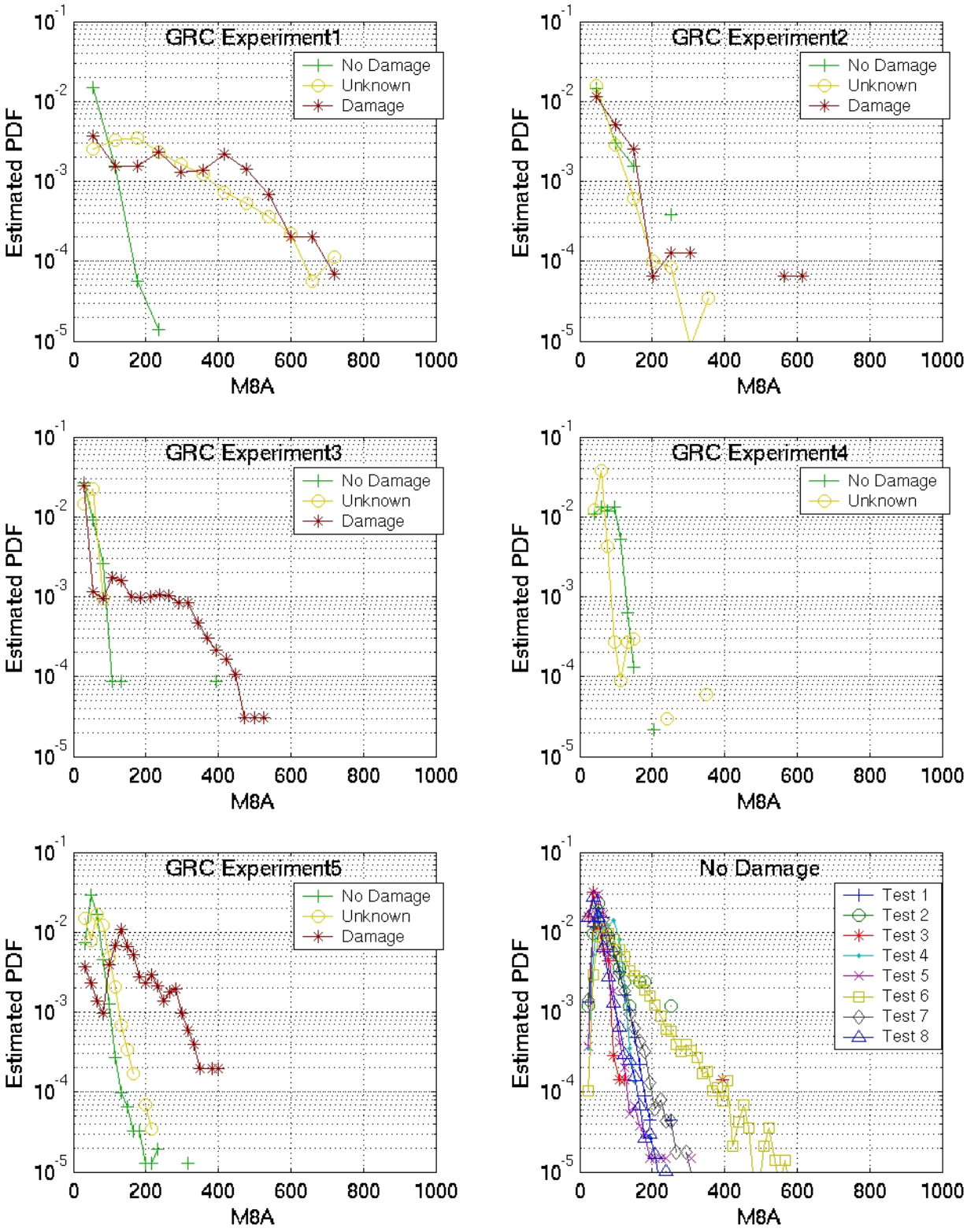

Figure 10. Estimated Probability Distribution Functions of M8A from GRC Test Rig Experiments

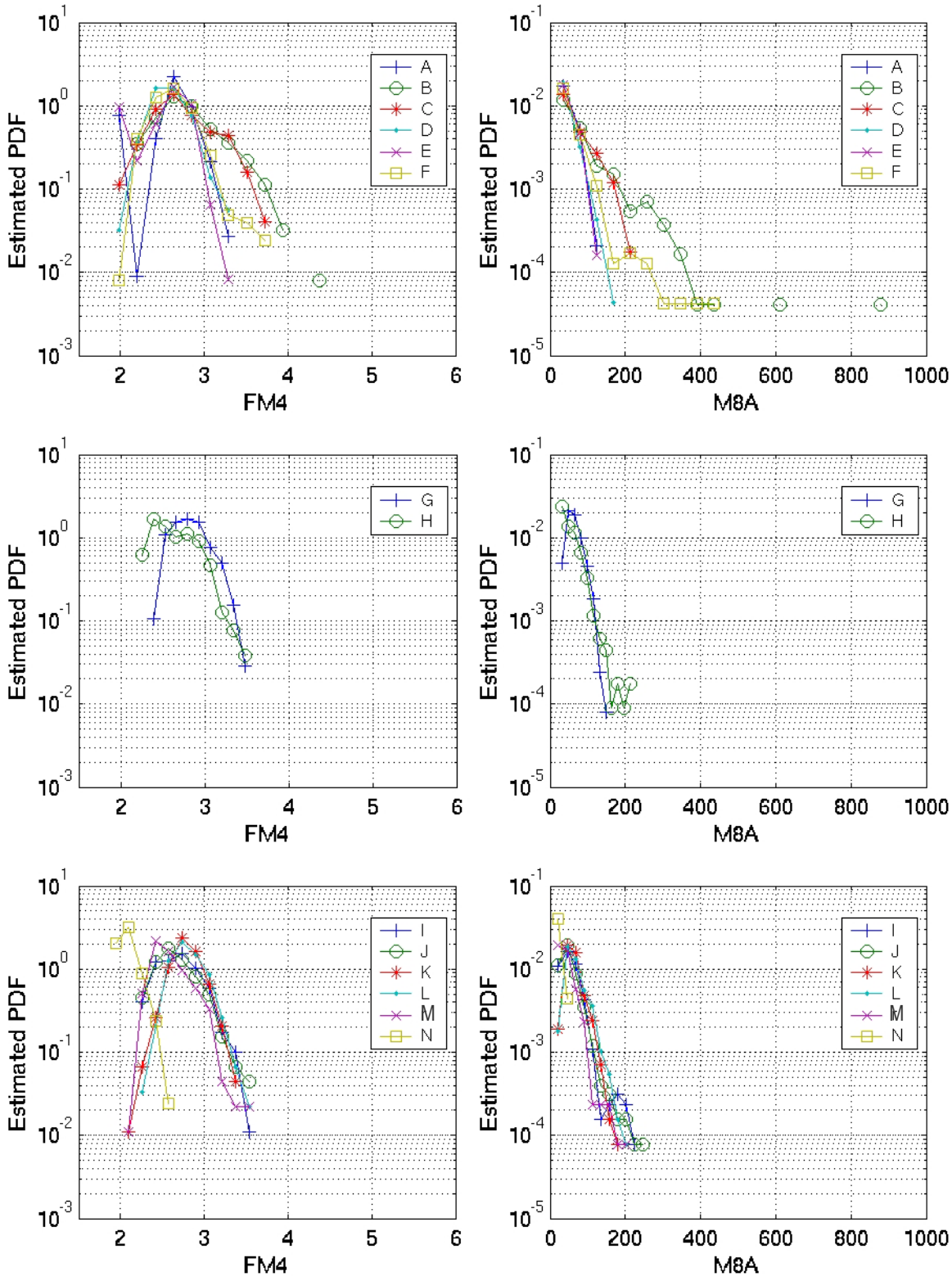

Figure 11. Estimated Probability Distribution Functions of FM4 and M8A from ARC Flight Test

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2003	Technical Memorandum

4. TITLE AND SUBTITLE

Threshold Assessment of Gear Diagnostic Tools on Flight and Test Rig Data

5. FUNDING NUMBERS

WBS–22–704–30–06

6. AUTHOR(S)

Paula J. Dempsey, Marianne Mosher, and Edward M. Huff

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
John H. Glenn Research Center at Lewis Field
Cleveland, Ohio 44135–3191

8. PERFORMING ORGANIZATION REPORT NUMBER

E–13812

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546–0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NASA TM—2003-212220

11. SUPPLEMENTARY NOTES

Prepared for the 59th Annual Forum and Technology Display sponsored by the American Helicopter Society, Phoenix, Arizona, May 6–8, 2003. Paula J. Dempsey, NASA Glenn Research Center; and Marianne Mosher and Edward M. Huff, NASA Ames Research Center. Responsible person, Paula J. Dempsey, organization code 5950, 216–433–3398.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category: 01 Distribution: Nonstandard

Available electronically at http://gltrs.grc.nasa.gov

This publication is available from the NASA Center for AeroSpace Information, 301–621–0390.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 words)***

A method for defining thresholds for vibration-based algorithms that provides the minimum number of false alarms while maintaining sensitivity to gear damage was developed. This analysis focused on two vibration based gear damage detection algorithms, FM4 and MSA. This method was developed using vibration data collected during surface fatigue tests performed in a spur gearbox rig. The thresholds were defined based on damage progression during tests with damage. The thresholds false alarm rates were then evaluated on spur gear tests without damage. Next, the same thresholds were applied to flight data from an OH-58 helicopter transmission. Results showed that thresholds defined in test rigs can be used to define thresholds in flight to correctly classify the transmission operation as normal.

14. SUBJECT TERMS

Gears; Transmissions; Oil debris sensor; Damage assessment; Damage detection; Health monitoring; Pitting fatigue

15. NUMBER OF PAGES

26

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	